Genetic Characterization of New Candidate Traits Derived from Test-Day Somatic Cell Counts

Jorge I. Urioste[1], Jessica Franzén[1] and Erling Strandberg[1]

Introduction

Mastitis is an inflammation in the mammary gland associated with adverse economic effects in dairy cattle. Genetic variation in susceptibility to the disease exists. However, routine recording of clinical mastitis (**CM**) is not performed in most countries; in many cases, measures of somatic cell counts (**SCC**) have been used in genetic evaluations, relying on the genetic correlation between these traits and mastitis (Mark *et al*. 2002; Miglior *et al*. 2005). Both mastitis and SCC have complex biological backgrounds (Harmon 1994). Despite of this, the lactation average of (the log of) SCC is often used as the only indicator of the former. A shortcoming of this measure is that the dynamic nature of mastitis is ignored. Alternative approaches have been explored (Detilleux and Leroy 2000; de Haas *et al*. 2004; Green *et al*. 2004; Madsen *et al*. 2008). Our scientific interest focused on novel traits that could be derived from useful information already present in the distribution of SCC curves. The new traits should be flexible enough to accommodate for sudden and drastic changes in SCC, especially those resulting from a clinical mastitis case; they should be considered as indicator traits for the economically important mastitis, and be simple to use in routine genetic evaluations. The objectives of this study were: (i) to define new traits that capture changes along the SCC curve; and (ii) to identify genetic variation in these traits.

Material and methods

Data. Data were obtained from Jälla research herd (Department of Animal Breeding and Genetics, SLU), between 1989 and 2004. Cows were of the Swedish Red (SR) and Swedish Holstein (SH) breeds, which were milked twice daily. Cases of CM were detected by the milkers and diagnosed by a veterinarian. Two datasets were created: Dataset **W** (weekly measures) included 1006 cow-lactation records, obtained from 467 cows, sired by 147 bulls (in total 36,051 TD observations). Dataset **M** (monthly measures) mimicked a field data set through keeping one observation from the complete dataset approximately every 30 days. It contained 980 cow-lactation records originated from 9159 TD observations from 457 cows sired by 147 bulls.

Traits. CM was scored as present (1) in a given lactation if at least one case of CM was recorded, and absent (0) otherwise. Twenty different alternative SCC traits were defined and designed to capture SCC general levels (total amount of SCC in milk, average test day SCC in early and late lactation), variation along the curve (standard deviation log SCC during lactation, Green *et al*. 2004), level of infection (presence and number of TD between set levels of SCC; number of peaks), and time of recovery (number of days sick in the widest

[1] *Dept. Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden*

infection peak, sum of all days along the peaks, or average number of days). A grid of arbitrary threshold values (40, 80, 150 and 500 thousand cells per mL) was applied for test-day SCC. Each test-day SCC was then classified individually as a binary trait between two any consecutive thresholds. An infection peak was defined as a period of increased SCC (over 150,000 cells/mL) between two low (<150000 cells/mL) TD observations. Patterns of SCC peaks could be also distinguished, based on number of days of increased SCC during lactation. Thus, length of sickness was recorded as the number of days between the start and the end of an infection peak (cell counts >150,000 cells/mL).

To select candidate traits among all possible traits, clustering procedures, canonical discriminant analyses and logistic regression with the logit of CM as response variable and the candidate SCC traits as explanatory variables in a stepwise procedure, were used. Traits that showed closest association with CM in the phenotypic analyses were subsequently examined for their genetic background. A list of the final traits considered in this study is presented in Table 1; two very often used production traits, Average test day milk and Average fat/protein ratio were also used as reference traits.

Table 1: Definition of the traits analyzed

Trait	Definition
CM	Susceptibility to clinical mastitis (0/1)
P>500	Test day SCC>500,000 cell/mL (0/1)
LogSCCSD	Log of SCC standard deviation
LogDSWP	Log of days sick in the widest peak
AVMILK	Average test day milk
AVFPR	Average test day fat/protein ratio

Methods. Single trait analyses were run for the candidate traits identified in earlier steps, using the following general mixed animal model: y = birth-year group (3 levels: 1985 to 1990, 1991 to 1995 and later than 1995) + breed (3 levels: Swedish Holstein; Swedish Red, high or low fat line) + lactation number (5 levels, first to fifth or more) + animal (1717 levels) + permanent effects in repeated cow measurements (467/457 cows) + residual

A threshold liability approach (e.g. Gianola and Foulley 1983) was used for traits expressed as a discrete response. In two cases (LogSCCSD and LogDSWP), the variables were log-transformed to improve normality. Genetic parameters (heritability, repeatability) were drawn from the posterior distributions using a Bayesian approach and Gibbs sampling, as implemented in the program Thrgibbs1f90 by Shogo Tsuruta, the University of Georgia, Athens, USA. Based on visual inspection of trace plots, a chain of 120,000 iterations was run for each trait, with a burn-in of 20,000 rounds, keeping every 50th sample for inference of posterior features.

Results and discussion

Overall mastitis rate was 23%; 81% and 50% of lactations had at least one test day with SCC above 150,000 or 500,000 cell/mL, respectively. Infection peaks were detected in 849 lactations (84.4% of total), with 754 lactations having between one and five peaks, and 95

having six or more. For cows with infection peaks, mean total number of days sick was 109 days, with an average of 43.2 days per peak.

To simplify interpretation of the analyses, discriminant procedures and a stepwise logistic regression of the logit of CM identified a subset of variables that explained most of the differences between clusters or minimized the error term, respectively (results not shown). A measure of SCC curve variation, the SCC standard deviation, and discrete (0, 1) measures of SCC depicting level of infections (P>500) were chosen. A variable illustrating the dynamics of infection (Log of Days sick in the widest peak) was also selected (Table 1).

In spite of the limitations posed by the datasets available, evidence of additive genetic variability was found for all the novel SCC traits (Table 2). Posterior heritability values ranged from 0.10 to 0.13 in dataset W, and 0.12 to 0.14 in dataset M, being larger than that found for CM, 0.04 in both datasets. As expected, production traits showed higher heritability estimates, 0.23 for milk yield and 0.52-0.58 for fat/protein ratio. Uncertainty in all estimates was large, however.

Table 2: Posterior median (percentiles 5 and 95) values of heritabilities for dataset W (weekly measures) and dataset M (monthly measures)

	Dataset W	Dataset M
CM	0.04 (0.01-0.11)	0.04 (0.00-0.11)
P>500	0.12 (0.03-0.24)	0.12 (0.04-0.24)
LogSCCSD	0.10 (0.03-0.19)	0.13 (0.06-0.21)
LogDSWP	0.13 (0.04-0.22)	0.14 (0.08-0.21)
AVMILK	0.23 (0.09-0.37)	0.23 (0.11-0.36)
AVFPR	0.58 (0.45-0.70)	0.52 (0.38-0.64)

Somatic cell counts are a useful measure of udder health, because they are widely recorded and are genetically associated to CM. One of their potential shortcomings as indicator traits is that acute, short-lasting infections may be difficult to identify simply from increased mean SCC during lactation, since SCC is recorded at approximately monthly intervals. It is apparent that different mastitis pathogens elicit different SCC responses; in particular, CM associated with 'environmental' organisms, such as *E. coli* may result in a high SCC for a short period of time (de Haas *et al*. 2004). Use of mixture models (e.g. Detilleux and Leroy 2000; Madsen *et al*. 2008) are theoretically attractive but currently difficult to use in a practical setting for prediction of breeding values, due to computational limitations.

The experimental herd data analyzed in this paper have limitations in structure and size, but these disadvantages are counteracted by the more consistent definition of clinical mastitis and the more frequent SCC testing. A relevant finding in this study is that traits such as the SCC standard deviation and presence of TD with SCC>500,000 cells/mL were selected in first place when trying to explain phenotypic relationships with CM; both traits showed sizeable genetic variation, turning into good candidates for selection against mastitis. Presence of patterns of peaks in lactations has also been shown to be phenotypically informative for pathogen specific CM (De Haas *et al*. 2004). Further, studies by de Haas *et al*. (2003; 2008)

have shown that traits similar to those developed here are heritable (0.01 to 0.13) and show sizeable genetic correlations with CM (0.55 to 0.93), which in many cases were stronger than that between CM and average somatic cell score. The heritabilities estimated in this study were in the range of 0.10 to 0.13, which is fairly consistent with those estimates. Our estimates showed higher heritabilities for the new SCC traits, compared with estimates for CM (Table 2), indicating good scope for selection. However, the uncertainty in the estimates was fairly large, and genetic correlations could not be properly estimated. This supports the possibility of carrying these results on to national databases, for more accurate estimations of genetic parameters and adaptation of the findings to field conditions. A combination of several definitions of SCC (indicator traits) might be necessary to capture the full scope of mastitis complexity. Then, genetic selection on, say, an index with several SCC traits could be proposed, and progress might be obtained if the index is applied by national milk recording systems. Some additional information could be gained from the patterns of peaks for CM (De Haas *et al.* 2008). Future studies in this direction are warranted.

Conclusions

The importance of this research is that it suggests some alternative traits derived from already existing information in test-day recordings of SCC. We have found evidence of genetic variation in the novel traits. Besides being reasonably heritable, the new traits are simple to define from existing information, inexpensive and easy to analyze using ordinary mixed linear models. Making better use of that information may therefore improve genetic programs aimed to select for lower susceptibility to CM.

Acknowledgements

References

De Haas Y., Barkema, H.W., Schukken, Y.H. *et al*. (2003). *Anim. Sci.*, 77:187-195
De Haas, Y., Veerkamp, R. F., Barkema, H. W. *et al*. (2004). *J. Dairy Sci.*, 87:95-105.
De Haas Y., Ouweltjes W., ten Napel J. *et al*. (2008). *J. Dairy Sci.,* 91:2501-2511.
Detilleux, J., and Leroy, P. L. (2000). *J. Dairy Sci.,* 83:2341-2349.
Gianola, D., and Foulley, J.L. (1983). *Genet. Sel. Evol.*, 15: 201–224.
Green, M. J. , Green, L. E. , Schukken, Y. H. *et al*. (2004). *J. Dairy Sci.,* 87:1256-1264.
Harmon, R. J. (1994). *J. Dairy Sci.*, 77:2103-2113.
Mark, T., Fikse, W. F., Emanuelson, U., *et al*. (2002). *J. Dairy Sci.,* 85:2384–2392.
Miglior, F., Muir, B.L., and Van Doormaal, B. J. (2005). *J. Dairy Sci.*, 88: 1255–1263
Madsen, P., Shariati, M. M., and Ødegård, P. (2008). *J. Dairy Sci.,* 91:4355-4364.